JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1993, 26, 281–291 NUMBER 3 (FALL 1993)

EFFECTS OF INVENTED SPELLING AND DIRECT INSTRUCTION ON SPELLING PERFORMANCE OF SECOND-GRADE BOYS

MARIBETH GETTINGER

UNIVERSITY OF WISCONSIN—MADISON

Four second-grade boys, 2 rated by their classroom teacher as below average and 2 as above average in basic language skills, participated in a 16-week spelling investigation. The participants alternately received, in counterbalanced order, 5 weeks of an invented spelling approach that incorporated 15-min creative writing periods and 5 weeks of direct instruction that involved 15-min periods of guided practice on spelling word lists. At the end of 10 weeks, each condition was replicated for 3 additional weeks. Although direct instruction resulted in more targeted words spelled correctly, invented spelling resulted in more nontargeted words spelled correctly, higher preference ratings by children, and higher teacher ratings of the quality of 3 of the children's writing samples.

DESCRIPTORS: invented spelling, direct instruction, spelling performance

Various instructional procedures for and theoretical accounts of the acquisition of spelling skills have appeared in the literature during the last decade (see Brown, 1990, for a review of research on spelling). Many research-based principles of effective spelling instruction have emerged from one of two contrasting models for the development of spelling competence (Brown, 1990). The first is a two-stage model, which posits that children progress from a stage of being unable to spell to being able to spell correctly. The act of spelling is conceptualized as a procedural task in which students are required to focus on, discriminate, and consistently reproduce letters of words in proper order (Doyle, 1983). From a behavior-analytic perspective, spelling success is attributed primarily to the use of highly structured, teacher-controlled environmental contingencies designed to strengthen the accurate production of written words. Teaching methods known collectively as *direct instruction* derive from this model of academic competence and have been used successfully to enhance children's performance in several academic skill areas, including spelling (Good, 1983; Jenson, Sloane, & Young, 1988; Rosenshine, 1983; Sulzer-Azaroff & Mayer, 1986).

Direct instruction involves systematic presentation of instructional materials and differential responses to students' academic performance. For spelling instruction, features of direct instruction include (a) a test–study–test format that incorporates immediate corrective feedback, (b) error-correction procedures, (c) positive reinforcement for correct spelling, (d) modeling and imitation of correct spelling, and (d) systematic repeated practice to learn a set of words (Fitzsimmons & Loomer, 1980; Gettinger, 1984; Vallecorsa, Zigmond, & Henderson, 1985). Among the direct instruction components studied, the use of error-correction procedures has received consistent support in the literature of applied behavior analysis for enhancing spelling accuracy (Foxx & Jones, 1978; Gettinger, 1985; Matson, Esveldt-Dawson, & Kazdin, 1982; Ollendick, Matson, Esveldt-Dawson, & Shapiro, 1980).

A second and alternative view is that spelling competence progresses through multiple developmental stages. Based primarily on observations of children's spelling, researchers have identified five stages of spelling competence through which children progress in an orderly fashion (Gentry, 1982; Henderson & Templeton, 1986). In this model, spelling errors do not indicate lack of spelling competence, as with the two-stage model; rather, errors

Thanks to Karen Kraemer for her assistance in the preparation of this manuscript.

Requests for reprints should be sent to Maribeth Gettinger, Department of Educational Psychology, 1025 West Johnson Street, University of Wisconsin, Madison, Wisconsin 53706.

reflect the level of a child's emerging or developing spelling ability. According to this view, the act of spelling is conceptualized as a strategic task rather than a strictly procedural one (Doyle, 1983). Research shows that error making tends to be systematic, representing incomplete knowledge or inefficient algorithms rather than absence of spelling mastery (Gerber, 1984, 1988; Henderson & Beers, 1980). From a behavior-analytic perspective, spelling success is attributed to frequent exposure to language stimuli such as engagement in repeated reading and writing activities that are developmentally appropriate. Exposure to correctly spelled words may occasion a child's correct spelling of the words in his or her spontaneous writing. Through differential reinforcement of only correct spelling, inaccurate spellings eventually are replaced with accurate ones.

Proponents of the developmental model theorize that improvement in spelling over time occurs because students adopt more accurate response-generation strategies (Gentry, 1982). This explanation is based on observed improvement in children's spelling that occurs even without direct instruction. For example, a child may spell the word *from* as "frum." This error reflects the use of a phonetic strategy. The transition to spelling *from* correctly reflects the child's use of a more sophisticated combined phonetic and visual-matching strategy. The latter strategy is reinforced when the child is exposed to the word *from* in reading and writing activities. Interestingly, Nulman and Gerber (1984) observed similar improvement over time in the developmental stage or quality of children's spelling when they received systematic instruction and error correction. Thus, there is some evidence that students taught with direct instruction also develop accurate strategies, suggesting that external contingencies for spelling achievement may contribute to the progressive development of spelling strategies. Methods known as *invented spelling* originated from the developmental view. To date, the number of investigations documenting the effectiveness of invented spelling are limited; nonetheless, this approach is gaining popularity in many schools (Downing, Coughlin, & Rich, 1986).

Although invented spelling shares many effective teaching procedures with direct instruction (e.g., contingent positive reinforcement of correct spelling), there are both theoretical and procedural differences between the two approaches that can be described in behavior-analytic terms. One salient difference is the nature of the targeted response. Rather than provide direct exposure to and repeated practice with isolated words in lists, teachers using invented spelling methods arrange repeated opportunities for students to generate their own representations of words, correct or incorrect, in the context of a writing passage. Whereas direct instruction targets and reinforces accurate reproduction of letter sequences, invented spelling targets use of effective spelling strategies and generation of written content. For example, with invented spelling, children are instructed to think about the sounds in unknown words and invent a way to spell each. One strategy for inventing spellings is for children to match their invented letter representations of sounds to the way that similar sounds are spelled in known words or in words that appear in reading material. Invented spelling relies heavily on children's own discriminations between accurate and inaccurate spelling. Direct instruction, however, relies on accurate production of letter sequences and, when necessary, teacher-guided discrimination between errors and correct spelling.

Another difference relates to the nature of the discriminative stimuli for spelling. In direct instruction, students attempt to spell words in response to teacher-provided overt visual and/or verbal stimuli (i.e., written or spoken word cues). Words are typically presented in lists, independent of context. In contrast, the stimuli in invented spelling are contextual and more covert. They represent an amalgamation of the learner's language experience and natural exposure to words and orthographic rules. Targeted words are embedded in contexts (e.g., stories, personal experiences) that typically include many words children already know how to spell. The discriminative stimuli, in effect, intersperse unknown words with known words, a strategy that has been shown to have beneficial effects on the spelling of unknown words (Brown, 1990;

Neef, Iwata, & Page, 1977). In sum, the discriminative stimuli in direct instruction are isolated words presented orally and/or visually by the teacher; in invented spelling, the discriminative stimuli are words, both known and unknown, in a context that the child generates on the basis of personal language experience.

A final difference is the consequences provided for spelling errors. In direct instruction, errors result in explicit corrective feedback to focus the student's attention to the stimulus word and to facilitate discrimination between errors and correct spellings. Feedback is followed by students' correction of errors and repeated practice in producing the correct spelling. In invented spelling, however, feedback (exposure to correct spellings of words that occur in language activities, or the provision of a correct model) is provided to increase the students' ability to apply orthographic knowledge to their written work. Although there is neither explicit correction of errors nor repeated practice of correct spellings, invented spelling promotes dicrimination between correct and incorrect spellings as well as reinforcement of accurate spelling. Within a behavior-analytic framework, therefore, the nature of the discriminative stimuli, the target response, and the type of feedback all differ significantly between direct instruction and invented spelling.

Few studies have systematically evaluated the effectiveness of invented spelling relative to direct instruction. Most of the literature supporting invented spelling includes anecdotal reports of successful applications (Forester, 1980; Lancaster, Nelson, & Morris, 1982; Lehr, 1986). However, as Groff (1986) noted, research has yet not confirmed that invented spelling improves spelling ability when compared with more structured, teacher-directed word-study approaches. The purpose of the present study was to compare the effects of these two instructional approaches on the spelling and writing performance of 4 second-grade boys. Based on the procedural differences between the two approaches and the nature of targeted responses that are reinforced in each, it was predicted that invented spelling would result in better overall writing (higher quality of writing and stronger preference for writing activities), whereas direct instruction would yield higher spelling accuracy on targeted words. Furthermore, it was predicted that these effects would occur for both above-average and below-average spellers.

The literature of applied behavior analysis often lacks the systematic investigation of experimental procedures that are implemented and evaluated in the context of ongoing classroom instruction and routines (Fuqua & Schwade, 1986). According to Vallecorsa et al. (1985), classroom teachers often fail to incorporate empirically valid techniques for teaching spelling, even when they are familiar with effective research-based principles and techniques. One reason for this is that many experimental procedures are often conducted in a restricted format and rely on one-to-one methods that may be difficult to implement in an actual classroom setting. Furthermore, data on children's performance during their regular spelling or writing assignments are seldom available. In light of this apparent discrepancy between classroom practices and research-based procedures, the present study also sought to increase the social validity of the experimental treatments by implementing them in the context of ongoing classroom instruction.

METHOD

Participants

Four 7-year-old boys in one second-grade classroom participated in this study. Two boys were identified by their classroom teacher as having below-average basic language skills; 2 were identified as having average to above-average basic language skills. No formal test data were available, but a spelling and word-reading pretest of the 96 words targeted for instruction was administered by the experimenter to all participants 1 week prior to the initiation of the study. The pretest showed a marked difference in the spelling and word-reading accuracy between the below-average and above-average students. Spelling accuracy scores (number correct) were 1 and 5 for the below-average students (Gary and William, respectively) and 10 and 9 for the above-average students (David and Todd, respec-

tively); word-reading accuracy scores were 3 and 6 for Gary and William, respectively, and were 23 and 19 for David and Todd, respectively.

Setting and Materials

The study was conducted in the children's classroom over a 16-week period. The curriculum in the participating school incorporated an integrated language approach to reading and writing. Because words are typically encountered in sentence and story contexts, this approach assumes that increasing students' experience with prose material, through reading and writing activities, will simultaneously improve their spelling and pronunciation of individual words (Quandt, 1983). As part of the integrated language approach, classroom activities (e.g., reading, writing, special art projects) are centered around a common weekly theme (e.g., birds, planets). For each theme, the classroom teacher developed a set of six targeted words to which children were exposed during reading activities and that children were instructed to use during their journal-writing assignments. Each six-word set contained five thematic words (e.g., *farm, cow, chicken*) and one high-utility word (e.g., *and*). Thus, a total of 96 words (six words for each of 16 weeks) were targeted for instruction and assessment.

The classroom in which the study was conducted had four different learning centers to which children circulated during a 90-min independent work period in the morning, three times each week. One learning center was a writing center; the other three centers focused on math problem solving, oral reading, and a cooperative project (e.g., making a map). Four to 5 children participated at each learning center at one time. Learning center periods were scheduled such that only 1 study participant was at the writing center at a given time.

Experimental Design and Conditions

A crossover design was used to compare the effectiveness of invented spelling and direct instruction (Kazdin, 1980). One below-average and 1 above-average participant were assigned at random to either the invented spelling condition or the direct instruction condition for 5 consecutive weeks. At the end of 5 weeks, the conditions were reversed for another 5 consecutive weeks. At the end of the second 5 weeks, each spelling condition was replicated for each pair of children for an additional 3 consecutive weeks.

Invented spelling. Invented spelling was implemented during 15-min periods of creative writing about weekly theme-related topics. At the writing center, children were instructed to write in their notebooks using the six targeted words that were posted on a bulletin board. The integrated language approach provided exposure to targeted words during a variety of classroom activities each week, including art, music, math, and reading. Thus, children were given the directive to generate a writing sample ("Write whatever you want to") prompted by a classroom activity or their own experiences and ideas. Examples of the content of children's writing included math story problems for their peers to solve, descriptions of art projects, or summaries of stories. Children also received prompts from the experimenter to spell inventively (e.g., "Try to write letters for every sound in the word") and were praised for their invented spellings irrespective of accuracy. After each individualized writing activity, the experimenter immediately rewrote what the child had written directly on the child's paper, providing correct models for all misspelled words. Aside from providing a correct model, the experimenter gave no direct instruction in spelling. Children did not practice writing words in isolation, and spelling errors did not result in additional discrimination training or writing practice. In sum, the role of the experimenter was (a) to provide the stimulus cues for writing (e.g., highlight classroom activities about which to write), (b) to reinforce (praise) the content or ideas expressed in children's writing as well as their invented spellings of targeted words, and (c) to ensure exposure to accurate spelling by providing correct models in context.

Direct instruction. Direct instruction occurred during 15-min periods of teacher-directed instruction and practice in spelling the week's targeted

words. At the writing center, children were guided through a look–cover–write–check study routine (Horn, 1947) until they had written each word correctly two consecutive times. The experimenter showed the correct spelling of each word separately on an index card, pronounced the word, and spelled it orally. Subsequently, the word was covered and children were instructed to write it from memory. After each attempt, children were shown the correct spelling. When an error occurred, the experimenter circled the error part of the word in the child's version, rewrote the word correctly, and specifically pointed out the similarity and difference to the correct model. Following this error-correction procedure, children rewrote the word, copying the experimenter's model. Occasionally, the above-average participants reached mastery on targeted words before the end of the 15-min period. As time permitted, students were instructed to write sentences using the targeted words; however, spontaneous writing was not specifically reinforced. In sum, the role of the experimenter was (a) to provide the visual and verbal stimulus word cues, (b) to reinforce (praise) correct spelling of targeted words, (c) to focus the child's attention on misspelled words, and (d) to ensure repeated practice in producing the accurate spelling of each word from memory at least two consecutive times.

Measurement

Accuracy of targeted words. On Friday of each week, children were given a dictated-word test of the six weekly targeted words by their regular classroom teacher. In addition, children were allowed 10 min to produce a writing sample that incorporated the six words. The classroom teacher pronounced the words once before the children were directed to generate their writing samples. If a child failed to use a targeted word in his spontaneous writing, it was scored as incorrect. If less than 50% of multiple uses of a targeted word were correct spellings, the word was also scored as incorrect. For example, if a targeted word was used three times, and two instances were incorrect, the word was scored as an error. Each child received two weekly accuracy scores, each ranging from 0 to 6, that reflected the number of targeted words spelled correctly in a dictated-word format and in spontaneous writing.

Overall spelling accuracy. Spelling accuracy was determined on the basis of the percentage of total nontargeted words (i.e., all words excluding targeted words) spelled correctly in the spontaneous writing samples. The difficulty and number of words generated in spontaneous writing varied across children but remained relatively constant over time for each child, thus allowing within-child comparisons. If less than 50% of multiple uses of a word were spelled correctly in the spontaneous writing sample, the word was scored as incorrect. Each child received a weekly spelling accuracy score on nontargeted words produced in spontaneous writing (ranging from 0% to 100%).

Accuracy on weekly targeted words reflected students' mastery of words that were taught, but accuracy on nontargeted words may indicate the degree to which spelling skills generalized to other words that were part of their writing vocabulary. The first score can be viewed as measuring treatment effectiveness and the second as measuring generalization of treatment effects. All spelling tests and writing samples were scored independently by two raters (graduate students in educational psychology) for spelling accuracy. There was 100% agreement between raters.

Overall quality of writing. A 5-point rating scale (excellent = 5; above average = 4; average = 3; below average = 2; poor = 1) was used by an independent rater to evaluate the overall quality of content and ideas expressed in children's weekly writing samples relative to other second-grade children. The rater was another second-grade teacher in the same school who did not know the participants and was naive to the experimental condition. Children's writing samples were typed and distributed to the rater in random order every week. At the end of the 16-week intervention, a second rater evaluated 16 writing samples (four per participant) selected at random. The second rater was also a second-grade teacher in the same school. The two

Table 1
Spelling Accuracy, Writing Quality, and Preference Ratings Across Experimental Phases

	Experimental phase							
	Invented Spelling 1 (5 weeks)		Invented Spelling 2 (3 weeks)		Direct Instruction 1 (5 weeks)		Direct Instruction 2 (3 weeks)	
Participants[a]	*M*	Range	*M*	Range	*M*	Range	*M*	Range
Spelling accuracy on targeted words (lists)								
David	3.4	3–4	4.3	3–5	6.0	6–6	5.7	5–6
Gary	1.4	1–2	2.3	2–3	3.4	3–4	3.7	3–4
Todd	3.6	3–5	4.0	3–5	5.4	4–6	5.0	4–6
William	1.2	0–2	1.3	1–2	3.4	3–4	3.7	3–4
Spelling accuracy on targeted words (writing samples)								
David	3.2	3–4	4.0	3–5	6.0	6–6	5.3	5–6
Gary	1.2	0–2	2.0	1–3	3.2	2–4	3.3	3–4
Todd	3.4	3–4	3.7	3–4	5.2	4–6	5.0	4–6
William	1.4	1–2	1.7	1–2	3.2	2–4	3.3	3–4
Spelling accuracy on nontargeted words[b]								
David	62	52–71	70	67–72	47	41–55	60	58–63
Gary	30	18–37	40	35–41	23	17–32	35	30–39
Todd	60	53–61	70	68–73	44	36–52	53	50–63
William	26	22–31	37	32–41	19	15–22	24	21–31
Quality ratings (1–5)								
David	4.0	4–4	4.7	4–5	3.2	3–5	3.7	3–5
Gary	2.0	2–2	3.0	3–3	1.2	1–2	1.7	1–3
Todd	4.3	3–5	4.3	4–5	3.6	3–5	3.3	3–4
William	1.8	1–2	2.3	1–3	1.9	1–3	1.7	1–3
Preference ratings (1–4)								
David	3.6	3–4	4.0	4–4	1.2	1–2	1.7	1–2
Gary	1.2	1–2	2.3	1–3	1.0	1–1	1.0	1–1
Todd	3.2	3–4	3.7	3–4	1.6	1–2	1.7	1–2
William	1.9	1–3	2.0	2–2	1.2	1–2	1.3	1–2

[a] David and Todd, above average; Gary and William, below average.
[b] Percentage of nontargeted words spelled correctly in weekly writing samples, averaged across weeks.

teachers had the same quality ratings (100% agreement) for 14 samples (87.5% of the samples) and differed by only 1 rating point for the remaining two writing samples (12.5%). Thus, each child received a weekly writing score, ranging from 1 to 5, that reflected the quality of content, irrespective of spelling accuracy, relative to other second-grade children's writing.

Preference for writing. Following the dictated-word test and writing activity on Friday, children were asked to rate their overall enjoyment of the 15-min writing center periods for the week from 1 (not at all) to 4 (very much). Each child received a weekly preference score, ranging from 1 to 4, reflecting his overall preference for the writing center activities.

RESULTS

Table 1 presents the weekly accuracy scores, averaged within phases, for targeted words on dictated-word tests and in writing samples (possible

range, 0 to 6). For all participants, there was no difference between the dictated-word and spontaneous writing formats in the number of targeted words spelled correctly. Figure 1 shows the number of targeted words spelled correctly each week for participants over 16 weeks. Because there was no difference in accuracy between dictated-word tests and writing samples, each data point in Figure 1 represents the average of the two weekly accuracy scores on targeted words. As predicted, children spelled more of the six words correctly when they received direct instruction and practice on words in isolation than during the invented spelling condition.

Table 1 presents the accuracy scores, averaged within phases, on nontargeted words that children produced in their end-of-week writing samples (possible range, 0% to 100%). Because the writing samples focused on a common theme, there was some overlap (approximately 30%) in the nontargeted words generated in children's writings. Nonetheless, none of the students produced exactly the same words in their writing samples. The results, therefore, must be interpreted cautiously because the percentages reflect spelling for words of varying length and difficulty. Despite these measurement limitations, the data in Table 1 show slightly higher overall accuracy scores for children who engaged in frequent writing in the invented spelling condition. The effect was observed for both below- and above-average students.

Table 1 also presents the quality-of-writing ratings and preference-for-writing ratings averaged within phases for each participant. Although there was some variability across weeks within a single phase, on the average the writing samples produced by 3 of 4 participants (David, Gary, and Todd) during the invented spelling condition were rated more highly by teachers (range, 2.0 to 4.7) than were the samples produced during the direct instruction condition (range, 1.2 to 3.7). It should be noted that for 2 children (David and Gary), quality ratings increased from the first 5 weeks of each phase to the second 3 weeks (which occurred 5 weeks later), reflecting some improvement in writing over time for these participants. Interestingly, quality ratings during direct instruction showed a decline over time for Todd and William.

Finally, as shown in Table 1, children also rated their preference for the writing center activities more highly during invented spelling (range, 1.2 to 4.0) than during direct instruction (range, 1.0 to 1.7). This effect was not as strong for the two below-average participants (Gary and William).

DISCUSSION

This study compared the effects of invented spelling to direct instruction on children's spelling and writing. Although there are limitations to the generalizability of these results (discussed below), the findings document different benefits for each approach. First, the results lend preliminary support to the claims made by proponents of invented spelling that the approach engenders greater enjoyment of writing activities and higher expressive quality in written work than does direct instruction. In contrast, weekly accuracy on targeted words was consistently higher during direct instruction than during invented spelling.

The observed average difference in quality ratings was minimal (approximately 1 point); however, it is noteworthy that 92% of all writing samples from the invented spelling condition were rated more highly than those from the direct instruction condition. In addition, the limited range in possible quality ratings (1 to 5) may have restricted the magnitude of rating differences across conditions. Despite these limitations, the differences in writing quality and preference ratings may be attributed to the opportunities for engaging in writing activities and reinforcement contingencies that differed between the two approaches. During direct instruction, children had few, if any, opportunities for creative writing at the writing center. During invented spelling, children engaged exclusively in creative writing, and their writing was consistently reinforced. The experimenter observed that children responded negatively (e.g., verbal complaints) during direct instruction when they misspelled a word and were required to rewrite it, and they responded positively when they avoided the error-correction

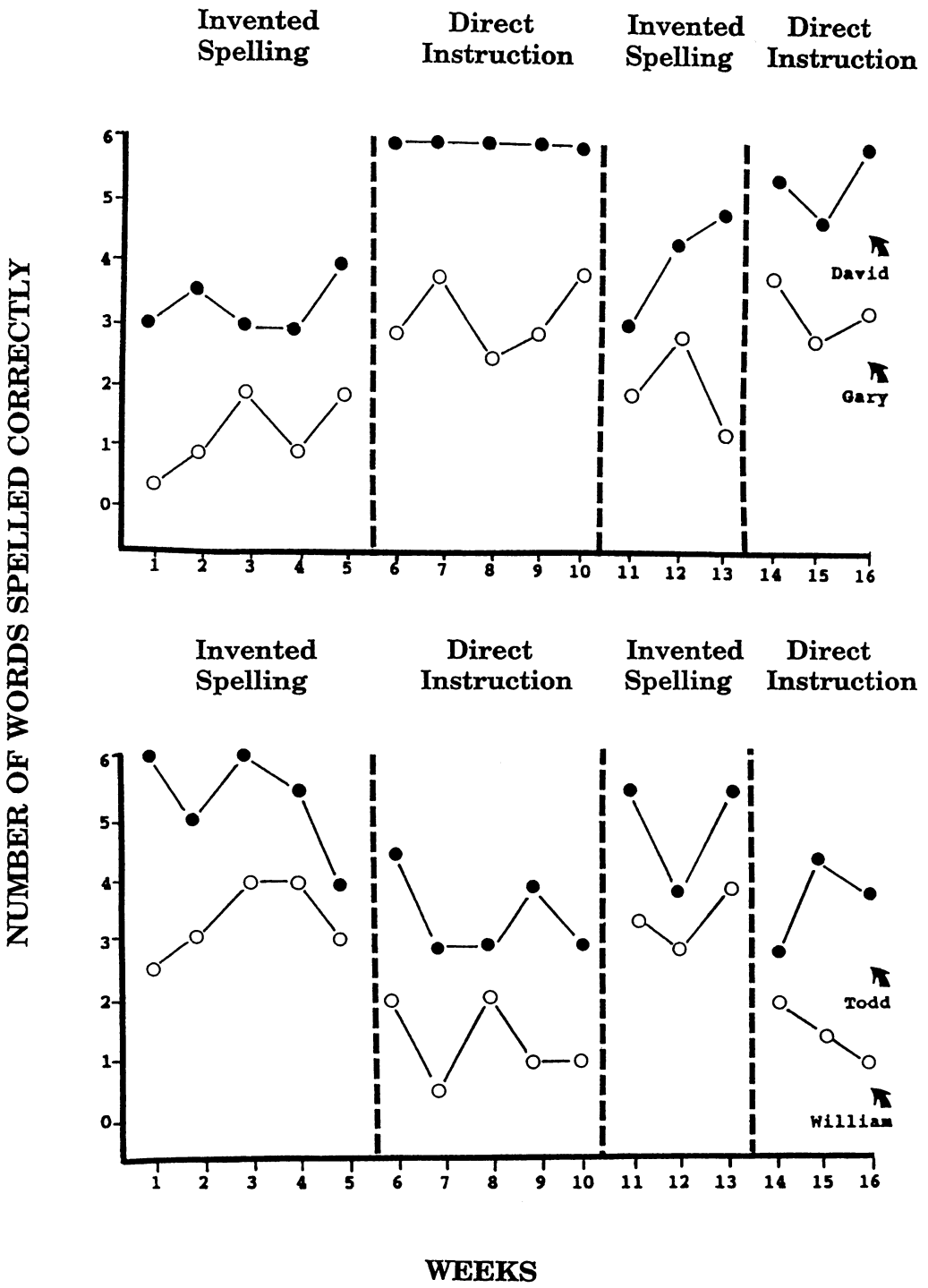

Figure 1. Number of targeted words spelled correctly each week across experimental phases (possible range, 0 to 6). Each data point represents an average of two weekly accuracy scores, one based on accuracy in dictated-word tests and the other on accuracy in writing samples.

and practice procedure by spelling the word correctly. In effect, all invented spellings (accurate as well as inaccurate) were reinforced during invented spelling, whereas inaccurate spellings were punished during direct instruction. This procedural difference may have led to a difference in preference ratings among students.

The effectiveness of each approach was gauged, in part, by students' accuracy in spelling each week's list of six words. In this case, as noted earlier, there is a clear advantage to direct instruction. It is interesting to note the absence of any apparent spillover effect from direct instruction to invented spelling phases. Children did not show any generalization of the study steps implemented during direct instruction (look–cover–write–check) to their learning of the targeted words during invented spelling. This is not altogether surprising. As part of the invented spelling approach, children were prompted to write words on the basis of the sounds they detected within each word, rather than the correct spelling (which was posted nearby). William, who received the direct instruction phase first, exhibited some reluctance to invent spellings during the first 2 weeks of invented spelling when asked to do so. He frequently requested help in spelling nontargeted words or produced a correct targeted word by copying from the posted list. His discomfort appeared to abate as the experimenter praised and encouraged his invented spellings.

Spelling accuracy on nontargeted words did not parallel the results for the targeted words. As shown in Table 1, children's spelling accuracy in the 10-min writing samples during invented spelling was equal to or slightly better than the accuracy in writing samples produced during direct instruction. At the same time, the invented spelling writing samples were rated more highly in content and ideas for 3 children, suggesting that, contrary to frequent criticisms of invented spelling, accuracy was not minimized in favor of more expressive quality in writing.

One plausible explanation for higher accuracy on nontargeted words is that each child appeared to have a repertoire of words in his writing vocabulary that were repeated frequently during all three weekly writing sessions of the invented spelling condition. For example, the theme for 1 week was birds. Although not on the list of targeted words, 1 child produced the word *worm* during each of his three weekly writing periods and again in his end-of-week writing sample. Invented spelling allowed him to attempt to spell the word *worm* three times and to see the correct spelling modeled by the experimenter after each attempt prior to the final writing sample. Inherent in invented spelling is this component of repeated opportunities to generate words and to view correct models. It appears that frequent opportunities to apply orthographic knowledge by generating invented spellings (correct or incorrect), even without directed word study and practice, had a positive effect on overall spelling accuracy. In fact, the percentage of targeted words spelled correctly was similar to the percentage of nontargeted words spelled correctly during the invented spelling condition for all participants (see Table 1 and Figure 1). In contrast, during direct instruction, the percentage of nontargeted words spelled correctly was lower than the high percentage of accurate spellings of targeted words. These results suggest that one of the major objections to invented spelling, that students may learn incorrect spellings if there is no direct instruction in how to spell words correctly, may not hold true, at least among beginning writers (Armington, 1984). An alternative or additional explanation for the higher accuracy on nontargeted words during invented spelling is that the response that is reinforced (i.e., orthographic problem solving) is more generalizable to spelling other words than is the response reinforced during direct instruction (i.e., accurate production of letter sequences from memory).

There are several factors that may affect the validity of these findings. First, the second-grade classroom in which the study was conducted incorporated an integrated language approach as part of its regular instruction for developing basic reading and writing skills. This orientation is more consistent with invented spelling than with direct instruction. Thus, the observed benefits of invented spelling may be linked to broader positive outcomes stemming from the overall instructional approach

adopted in the classroom. Although the instructional-setting variables were consistent across all participants and phases, the benefits may have been maximized for invented spelling because of its similarity to the integrated language orientation. A second limitation lies in the absence of spontaneous writing practice during direct instruction. Structured word study and practice replaced invented spelling and writing activities during this condition. It is possible that the higher quality ratings during invented spelling were functionally linked to more practice and higher engagement in writing activities rather than the reinforcement of children's writing content and invented spellings of words.

Another potential limitation stems from the use of the crossover design. Although the design attempted to control for sequence effects by counterbalancing replications, the integrity and independence of the two conditions may have been compromised, thus minimizing the observed differences in effects. For example, the improvement in quality-of-writing ratings over time for 2 participants (irrespective of treatment condition) suggests that the emphasis on writing during invented spelling may have carried over to the direct instruction condition. Finally, there are limits to the generalizability of the results given the limited number of participants, the relatively short intervention period for each treatment phase, and the focus on one grade level. The use of second-grade students may be especially problematic. Many of the goals of invented spelling, particularly those related to student control and motivation, are more salient among older students. Likewise, one objection to invented spelling, that children will learn that accuracy is not essential in spelling, is potentially stronger for older students than for second-grade students (Yule, 1986).

It may be that a combination of both approaches, invented spelling paired with systematic error correction and practice of misspelled words, will yield the greatest benefits for beginning spellers. As Doyle (1983) suggests, although students are naturally inventive in their spelling, teachers may need to provide additional task-specific or word-specific instruction, especially for low-achieving students. Reinforcement of invented spelling may need to be paired with direct teaching of basic spelling skills to facilitate both mastery of these skills (as demonstrated in this study) and the acquisition of orthographic knowledge and rules. Future research should continue to investigate the classroom application of invented spelling and direct instruction, particularly with entire classes of students, and to evaluate systematically the separate and combined effectiveness of both approaches.

REFERENCES

Armington, D. (1984). Invented spelling. *Teachers & Writers Magazine, 15,* 1–8.

Brown, A. S. (1990). A review of recent research on spelling. *Educational Psychology Review, 2,* 365–397.

Downing, J., Coughlin, R. M., & Rich, G. (1986). Children's invented spellings in the classroom. *The Elementary School Journal, 86,* 295–303.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53,* 159–199.

Fitzsimmons, R., & Loomer, B. M. (1980). *Spelling: The basics.* Iowa City, IA: University of Iowa Press.

Forester, A. D. (1980). Learning to spell by spelling. *Theory into Practice, 19,* 110–118.

Foxx, R. M., & Jones, J. R. (1978). A remediation program for increasing the spelling achievement of elementary and junior high school students. *Behavior Modification, 2,* 211–230.

Fuqua, R. W., & Schwade, J. (1986). Social validation of applied behavioral research: A selective review and critique. In A. Poling & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (pp. 265–292). New York: Plenum.

Gentry, J. R. (1982). Developmental spelling: Assessment. *Diagnostique, 8,* 52–61.

Gerber, M. M. (1984). Developmental aspects of learning to spell. *Academic Therapy, 20,* 11–19.

Gerber, M. M. (1988). Cognitive-behavioral training in the curriculum: Time, slow learners, and basic skills. In E. L. Meyen, G. A. Vergason, & R. J. Whelan (Eds.), *Effective instructional strategies for exceptional children* (pp. 45–64). Denver, CO: Love.

Gettinger, M. (1984). Applying learning principles to remedial spelling instruction. *Academic Therapy, 20,* 41–48.

Gettinger, M. (1985). Effects of teacher-directed versus student-directed instruction and cues versus no cues for improving spelling performance. *Journal of Applied Behavior Analysis, 18,* 167–171.

Good, T. L. (1983). Classroom research: A decade of progress. *Educational Psychologist, 18,* 127–144.

Groff, P. (1986). The implications of developmental spelling research: A dissenting view. *The Elementary School Journal, 86,* 317–323.

Henderson, E. H., & Beers, J. W. (Eds.). (1980). *Developmental and cognitive aspects of learning to spell—A reflection of word knowledge.* Newark, DE: International Reading Association.

Henderson, E. H., & Templeton, S. (1986). A developmental perspective of formal spelling instruction through alphabet, pattern, and meaning. *The Elementary School Journal, 86,* 305–316.

Horn, T. D. (1947). The effect of the corrected test on learning to spell. *The Elementary School Journal, 47,* 277–285.

Jenson, W. R., Sloane, H. N., & Young, K. R. (1988). *Applied behavior analysis in education: A structured teaching approach.* Englewood Cliffs, NJ: Prentice-Hall.

Kazdin, A. E. (1980). *Research design in clinical psychology.* New York: Harper and Row.

Lancaster, W., Nelson, L., & Morris, D. (1982). Invented spellings in Room 112: A writing program for low-reading second graders. *The Reading Teacher, 35,* 906–911.

Lehr, F. (1986). Invented spelling and language development. *The Reading Teacher, 39,* 452–455.

Matson, J. L., Esveldt-Dawson, K., & Kazdin, A. E. (1982). Treatment of spelling deficits in mentally retarded children. *Mental Retardation, 20,* 76–81.

Neef, N. A., Iwata, B. A., & Page, T. J. (1977). The effects of known-item interspersal on acquisition and retention of spelling and sight reading words. *Journal of Applied Behavior Analysis, 10,* 738.

Nulman, J. H., & Gerber, M. M. (1984). Improving spelling performance by imitating a child's errors. *Journal of Learning Disabilities, 17,* 328–333.

Ollendick, T. H., Matson, J. L., Esveldt-Dawson, K., & Shapiro, E. S. (1980). Increasing spelling achievement: An analysis of treatment procedures utilizing an alternating treatments design. *Journal of Applied Behavior Analysis, 13,* 645–654.

Quandt, I. J. (1983). *Language arts for the child.* Englewood Cliffs, NJ: Prentice-Hall.

Rosenshine, B. (1983). Teaching functions in instructional programs. *The Elementary School Journal, 83,* 335–351.

Sulzer-Azaroff, B., & Mayer, G. R. (1986). *Achieving educational excellence: Using behavioral strategies.* New York: Holt, Rinehart, & Winston.

Vallecorsa, A. L., Zigmond, N., & Henderson, L. M. (1985). Spelling instruction in special education classrooms: A survey of practices. *Exceptional Children, 52,* 19–24.

Yule, V. (1986). The design of spelling to match needs and abilities. *Harvard Educational Review, 56,* 278–297.

Received July 26, 1992
Initial editorial decision October 15, 1992
Revisions received January 14, 1993; March 31, 1993; May 7, 1993
Final acceptance May 7, 1993
Action Editor, F. Charles Mace